PTO/PCT Rec'd 3 DEC 2001 09/914400

ENDOSCOPIC OBSERVATION DEVICE

The present invention concerns a device for endoscopic observation of a field in three dimensions (3D) intended, in particular, but not exclusively, for surgery.

The purpose of the present invention is determining the distance separating a surgical tool from the various elements constituting the operating zone at high video frequency (that is in more than 25 Hz), and thus provide to the surgeon spatial data necessary for a good execution of operating movements.

Endoscopic approach is a surgical technique that acquires more importance daily. The principle of this approach, referred to as a "minimally invasive" approach, consists in practicing short incisions through which, by way of an adapted device, one penetrates the tools into the patient's body which are necessary for the intervention, as well as a camera or endoscope supplying on a screen images which permit the vision of the operative field. This technique has such a potential to reduce operating trauma and, at the same time, the costs associated to the hospitable stay, that all the sectors of the surgery are interested in it: digestive surgery, gynecology, ORL, cardiology, osseous surgery.

One of the greatest problems associated with endoscopic techniques is that working environment is shown in two dimensions (2D) while the operating movements occur in a three-dimensional (3D) space. Several attempts were made previously to acquire the three-dimensional information necessary for the execution of a surgical procedure and to provide that information to the surgeon in real time (e.g., at least at 25 Hz).

A first known approach consists in using mechanical arms at the end of which tools are fixed. This approach is penalized by a significant limitation of authorized movements and an important operating congestion of the zone, in particular when several tools are actuated simultaneously.

A second approach is the optical tracking that consisting of fixing to the instruments active elements (LED) that are monitored by cameras. However, the presence of cables, angular inflexibility and problems of sterilization made this method little used.

A third approach is the passive tracking that consists in mounting marks (generally, three spheres) on the tools that are tracked by a vision system. The problems of congestion are certainly reduced, but they remain. Besides, it is not possible to eliminate completely the risk of losing information about position because of an obstruction "cameras - marks" of the field of vision following, for example, a movement of the surgeon in the operating zone.

Inspired by the developments of interfaces used in virtual reality to be able to show in 3D info-graphic environments, new approaches were proposed. Using various types of glasses (passive or active), they exploit the capacity of the human brain reconstitute the notion of depth by presenting images with a parallax for the right eye different from that for the left eye. Problems of discomfort associated with wearing the glasses, the potential interruption of the signal of synchronization in the case of the active glasses and, especially, the restoration of a subjective image limited, in that case also, the use of such approaches.

Another similar device is described in the U.S. patent No. 4,935,810. It contains two view recording devices and two screens on which appear the respective images of the view recording devices. By selecting particular points on screens, it is possible to calculate the distance which separates them. Such a solution permits the knowledge of the distance which separates an object in an operating field, for example, a bistoury, from an organ to be treated. Regrettably, as for view recording operations, since the selection of the considered points and the calculation of distances is carried out in a sequential way it is then not possible to follow permanently the operative field. Not unimportant risks of manipulation errors result from this aspect.

The present invention provides a device for endoscopic observation free from the inconveniences and limitations of state-of-the-art devices.

In a more precise way, the device according to the invention comprises

at least two video cameras (18a , 18b) each associated with optics for acquiring images of a common area of the field to be observed and delivering an electric signal representative of the acquired images;

an electronic device (20) for processing the signals delivered by the two cameras;

means for placing in memory representations of objects that appear in said field;

tracking means for automatically identifying at least two points P_1 and P_2 that are common to the acquired images, wherein at least one point is associated with one of said representations, and wherein the tracking means produces information relative to the position of these points in three-dimensional space;

calculation means (26c) to determine, from said information, a value representing the distance separating said points;

processing means (32) to transform said representative value into signals; and

communication means (22) to provide the operator with the information relative to this distance from signals coming from the processing means.

Preferably, the two optics are rigidly associated one to each other to form a stereoscopic endoscope. The two optics further have parallel axes, spaced one from each other at a distance D; have equal focal lengths, f; and have coplanar focal planes. The data produced by the tracking means is then, comprises, for each of the two points P_1 and P_2, coordinates xL, xR and y of the image in the focal plan of the corresponding optics, where xL is the image x-coordinate of the point in the left space, xR is the image x-coordinate of the point in the right space and y is the y-coordinate of the point in the left and right spaces.

The calculation means (264) perform the steps of calculating the coordinates Xp_1, Yp_1 and Zp_1 of the point P_1, Xp_2 and Yp_2 and Zp_2 of the point P_2 in a cyclopean space using the formulae:

$$X = (D / \Delta).\ (xL + xR),$$

$$Y = D.\ y / \Delta, \text{ and}$$

$$Z = -\ D.\ f / \Delta, \text{ wherein } \Delta = xL - xR;$$

calculating the differences $Xp_1 - Xp_2$, $Yp_1 - Yp_2$, and $Zp_1 - Zp_2$; and

determining the distance separating the two points P_1 and P_2 in three-dimensional space using the formula:

$$d_{12} = [(Xp_1 - Xp_2)^2 + (Yp_1 - Yp_2)^2 + (Zp_1 - Zp_2)^2]^{1/2}.$$

According to a preferred embodiment, the device according to the invention furthermore presents the following features:

communication means comprising a video screen (22) displaying an image of the observed field;

processing means (32) generating a significant image of said representative value and superimposing the significant image on the image of the observed field on the video screen;

This significant image can be, among numerous embodiments, an index or a zone showing a color gradient.

The device further comprises means (18c) for synchronizing the images of the two cameras, an analog to digital converter (18d, 18-e) interposed between each camera (18a, 18b) and the respective electronic device (20) for processing the signals, a memory for the storage of signals coming from the analog to digital converters, a means for selecting (30a) only a part of the information coming from at least one of the converters, in order to reduce the volume of processed information, correction means (30b) to process the images in order to reduce the effect of any aberrations of the cameras, and switch (23) that can limit the information displayed on the screen to the image coming from one of the cameras.

Brief Description of the Drawings

Other features of the invention will appear from the following description with comparison to the appended drawings, in which:

Figure 1 illustrates, in broad outline, an assembly containing a device according to the invention applied to a surgical operation;

Figure 2 presents the general structure of the device according to the invention;

Figure 3 shows, in a more detailed way, a part of the structure of Figure 2;

Figure 4 illustrates the principle of calculation of the distance between two points in a three-dimensional space;

Figures 5a, 6a, 7a and 8a represent diagrams of classes, respectively the main diagram of the application and the diagrams relative to the acquisition, stereo and the tracking of tools, while figures 5b, 6b, 7b and 8b are diagrams of

sequences corresponding to the diagrams of classes containing the same number.

Figure 1 shows, in a very simplified way, the implemented means, according to the invention, for a surgical operation using a endoscopic observation device. The operation takes place in an operative field **10** internal in a body and permits intervention on an organ **11** by means of a tool **12** introduced into the body by an incision **14a**. The surface of the organ **11** is provided with marks **11a** that are, for example, points made by means of a biocompatible ink, pastilles members, or certain parts of the organ itself that have a particular appearance. Marks **12a**, generally formed with pastilles of color, are advantageously arranged on the tool **12**, so as to facilitate the identification and, as it is explained further below, calculating position of the tool.

In the device according to the invention, the observation of the operative field **10** is carried out by means of a conventional endoscope **16** with double optics, associated to a light source (not represented) and introduced into the body through another incision **14b**. The endoscope **16** is connected with a image capture device **18** containing two cameras **18a** and **18b**, called respectively left camera L and right camera R, which receive bright radiation picked up with the two optics, so as to be able to process images in a stereoscopic way, explained further below.

Cameras **18a** and **18b** transform, in a classic way, bright radiation resulting from the operative field **10** into an electric signal that is applied, by two different ways L and R, to processing electronics **20**. A video screen **22**, acting as communication device, displays with a frequency of 25 Hz images received with cameras **18a** and **18b**, before or after processing by the electronics **20**. A switch **23** permits the selection of one or other one of the cameras **18a** and **18b** or another image obtained after treatment by the electronics **20**.

As described further, the device according to the invention permits the definition of the distance between two marks **11a** and/or **12a** visible in the operative field **10**, but also between any points recognizable by their shape, their color, etc.

Referring now to Figure 2 that represents the general structure of the device according to the invention, one finds the operative field **10** with the organ to be treated **11**, the tool of intervention **12**, the endoscope **16**, the image capture device **18**, as well

as the various modules constituting the processing electronics **20**. The field **10** is fixed to a Cartesian frame, of which the Z axis is parallel to the axes of the cameras.

The image capture device **18** contains, besides the left **18a** and right **18b** cameras, a image synchronization circuit **18c** and two analog to digital (A/D) converters **18d** and **18e**.

The synchronization circuit **18c** permits the handling of images resulting from the two cameras in a way that they are perfectly synchronous, that facilitates their comparison, as is explained further, and improves the quality of information relative to the third dimension.

The two analog to digital (A/D) converters **18d** and **18e** transform analog signals coming from the synchronization circuit of **18c** into digital signals.

The output of the device **18** is connected to the input of a digital video recorder **24** which can perform the recording of all or any of the operation. In an alternative embodiment, not represented in the drawing, it is also possible to use an analog recorder which is then connected to the input of A/D converters **18d** and **18e**.

The output of the device **18** is also applied to the processing electronics **20**, advantageously comprising a computer, in which the various modules are defined by a set of programs and subroutines, described in reference to Figures 5 - 8. This computer contains conventional means of command not represented in the drawing, such as a keyboard and/or a mouse.

In an alternative embodiment, electronics **20** could be realized by means of various electronic modules. However, this second solution offers less flexibility of use.

To facilitate understanding, the device according to the invention is described, in reference to Figures 2 and 3, on the basis of a modular construction in which each module performs a particular function. Figures 5 - 8 will then allow better understanding of the structure of the software controlling the device.

The processing electronics **20** is ordered so as to process in parallel signals coming from the two cameras. However, to simplify the drawing, each of the modules or systems is represented only once.

The heart of the electronics **20** is constituted of a signal treatment module **26**, which includes object tracking systems **26a** and **26b**, intended to follow the possible

movements of objects (tools and organs) to which marks **11a** and **12a** are connected, as well as an unity of calculation **26c** of the distance separating at least two points associated with these marks in a three dimensional space. These last ones are selected, for example, by means of the computer mouse, among a set of representations put in memory of objects that may appear in the operative field. They are identified by the object tracking systems **26a** and **26b**.

The module **26** allows, as such, the determination of the position in the space of points of interests identified by the operator with the marks **11a** and **12a**, as well as the distance separating them. However, in order to confer on the device a greater flexibility of use and an optimal precision, electronics **20** furthermore contains:

an image acquisition module **28** containing an interface **28a** and a memory **28b**;

a correction module **30**, comprising a filtering circuit **30a** receiving signals from at least one of the A/D converters **18d** and **18e**, in order to select a part of the received signals, and a circuit **30b** for correction of the aberrations of cameras; and

a screen control module **32**, which comprises a system **32a** for combining images and a screen command module **32b**.

The acquisition module **28** is connected by its input to the output of the image capture device 18. The acquisition module **28** allows the storage of information and can thus play the role of black box, recording in its memory **28b**, the contents of which cannot be corrupted, all or part of the information relative to the progress of the operation. Its interface **28a** allows the transformation of the information received into a form compatible with the characteristics of the memory **28b**.

The correction module **30** is connected, by its input to the output of the acquisition module **28**. Its filtering circuit **30a** permits the processing of images to make them more legible. By applying them an adequate algorithm, one can keep only the outline of the objects appearing in the field **10** or, still, represent only one color of the images, strengthen the contrast, etc. This processing allows different view of the operative field and thus better capture certain details. As for the correction circuit **30b**, it functions to correct the aberrations of the optics of the endoscope. This correction is indispensable to be able to have a good image. Indeed, the optics of endoscopes are of

very small dimensions. It results a strong distortion of images. To avoid this inconvenience, one establishes a plane transformation enabling finding correspondences between the points of an idealized image and a deformed image. This method is described in a detailed way in Wolberg, George, *Digital Image Warping*, IEEE Computer Society Press Monograph, 1994.

Signals available at the output of the module **30** thus show characteristics that enable the display of information on the video screen **22** that is more easily interpreted by the surgeon. These signals are introduced into the treatment module **26** that is described in a more detailed way with reference to Figure 3.

Finally, the control module **32** performs the management of information displayed on the video screen **22**, combining or not the images of the operative field **10** with information concerning the position of the various present objects in the field **10**. The image combination system **32**a, is called “multimage”, by a contraction of “multiple” and “image”, or in English, “overlay.” It has its input connected to the output of module **26** and processes signals supplied by this module simultaneously with signals produced by the camera **18a** or **18b**. The screen command **32b**, connected to the output of the system **32a**, is connected by its output to the video screen **22** which permits the display of the operative field **10** as well as information relative to organs and to tools, in particular information relative to distances.

The indication of distances can be made in various manners. The value of the coordinate Z of the distance between the tool **12** and a selected mark **11a** can be displayed in digital format or, still, for example, by means of two indexes in V. It is also possible to represent the distance between a given point, for example, the end of the tool **12**, and all or part of the field **10** by a gradient of colors, for example, blue corresponding to the very remote zones and red those in contact.

Since the tool **12** has to approach with precaution to a particular point, this point being identified, it is also possible make a part of the tool 12 appear in a color corresponding to measured distance.

The choice of one or the other solution, practicable at any time, is made according to the problems with which the surgeon is confronted, giving more or less of

importance for information relative to the value of component Z or, on the contrary, to the operative field.

It is also possible to add additional images, for example, reticules, masks or any other image that may facilitate the work of the surgeon, that come from a data base not represented on the drawing and facilitate a diagnosis. The way of proceeding to obtain this superposition of images is clarified in reference to Figure 5.

To determine the distance between two marks belonging to objects arranged in the field **10**, it is necessary to establish a correlation between the various objects visible on the two images captured with cameras **18a** and **18b**. As shots are made from two different points, it results that the distances between the projections of the two marks in the planes of the cameras, transposed into a common frame of reference, are different. By calculating the difference of the coordinates of these projections in the common frame of reference, it is possible to calculate the distance separating these marks in the three dimensional space.

Figure 3 permits a better understanding of the way to proceed. It represents the detail of the unity of calculation [calculation unit?] **26c** which contains, connected in series, a geometrical correction entity **260** adapted to process images in epipolar geometry, an outline detection entity **262**, an correlation entity **264**, a distance determination entity **266** and a filter **268**.

The geometrical correction entity **260** allows to process images in epipolar geometry. To well understand this geometry, one will refer advantageously to the Zhang, Zhengyou, Determining the Epipolar Geometry and its Uncertainty: A Review, J. Computer Vision, 1998.

The outline detection entity **262** permits selecting, identifying and following marks **11a** and **12a**, as well as particular zones of objects being in the field **10**. The device can thus recognize the various objects present in the field and follow them in successive images. Means implemented to make this recognition are described in a complete way in Konolige, K.," Small Vision: Hardware Implementation ", Eighth International Symposium on Robotics Research, Hayama, Japan, 1997 ".

From information obtained by the outline detection entity **262**, it is possible to examine the two same points on both images, to define the distance which separates

them on both of these images and to determine the difference between these two distances. This difference, called disparity, is defined by the correlation entity **264**.

By knowing the characteristics of optics and the disparity, it is possible to establish a relation between the disparity and the distance between both points. This operation is made in the distance determination entity **266** and is described more in detail in reference to Figure 4.

Finally, the filter **268** performs the elimination of inaccuracies, for example by applying a method of spatial and temporal interpolation, such as that defined in the research paper 2013 of the INRIA (1993) entitled "Real-time correlation-based stereo: algorithm, implementations and applications", by Olivier Faugeras.

Figure 4 permits an understanding of the way of determining the coordinates of two points common to the images captured with the cameras, then calculating the distance that separates them. This figure shows two optics L (left) and R (right), having the same focal length f, and parallel optical axes at a distance of value D. Furthermore, their focal planes are coplanar. One sees, besides, a point P belonging to the field **10** and for which the position has to be determined.

For that purpose, one have two plane frames of reference, one left and one right, of which the respective x-coordinates, aligned on the same axis, are xL and xR and of which the y-coordinate is y. The distance between the origins of both frames of reference is equal to D. They allow respectively to define position pL of the image of the point P seen with the left camera and its position pR seen with the right camera. It is thus possible to measure the coordinate of the images of the point P in left and right frames of reference and, in doing so, to define its coordinates xLp, xRp and yp.

One can notice that the determination of coordinates in the left and the right frames is simply made by identifying the pixels of the cameras on which the image of the considered points are formed. The units of xL, xR and y are therefore pixels. Focal length f must be also expressed in this unit.

To define the position of the point P in the field, one uses a third frame of reference in which the axes X, Y and Z define a space called Cyclopean. Plane X-Y is parallel to plane xL/xR-y arranged in the front, at a distance equal to f. Axis Z is parallel to the optical axes and arranged in the same plane, in median position.

To calculate the coordinates of the point P in the Cyclopean space, one begins by defining disparity Δ of its images according to the formula:

$$\Delta = xL - xR$$

One can then define the coordinates of the point P in the Cyclopean space according to formulae:

$$Xp = (D / 2\Delta).(xL + xR)$$

$$Yp = D.y/\Delta$$

$$Zp = -D.f/\Delta$$

When the distance between two points P_1 and P_2 has to be calculated, one has thus to determine, first of all, according to the method above, the coordinates Xp_1, Yp_1 and Zp_1 of the point P_1 and the coordinates Xp_2, Yp_2 and Zp_2 of the point P_2. Distance d_{12} which separates them is then obtained with the formula:

$$D_{12} = [(Xp_1 - Xp_2)^2 + (Yp_1 - Yp_2)^2 + (Zp_1 - Zp_2)^2]^{1/2}$$

If, in formulae above, D and expressed in mm, the same will apply to values of X, Y and Z.

In the device according to the invention which has just been described, cameras **18a** and **18b** permanently perform the acquisition of the images of the operative field **10**, through the endoscope **16.** These images are synchronized with the circuit **18c** and converted from the analog mode to the digital mode by converters **18d** and **18e**. Signals are then sent to the recorder **24** in order to store the operation and to the image acquisition module **28** that permits, on the one hand, their modification by means of the interface **28a** by acting, for example, on the contrast, the luminosity, etc. and, on the other hand, to place them in non-volatile memory **28b**, being able to hold place of black box.

Signals thus obtained are applied to the correction module **30** that implements the elimination of certain defects affecting the quality of images. They are then sent to the processing module **26** that calculates the distances and the tracking of organs, then in the control module **32** that implements the superimposition of images before controlling the video screen **22**.

As explained above, the essential functions of the device according to the invention can be performed by software, advantageously written in an object-oriented

language. This software is outlined using a diagram of classes in Figure 5a and a diagram of the sequences of the main loop in Figure 5b. These diagrams refer to the analysis by the methodology UML described in Lai, Michel, "UML, Unified Notation Of Object Modelization, Application In Java", InterEditions ,1997, ISBN 2-7296-0659-9 and performed using the "Rational Rose" software, Rational Software firm.

In the following description, one can call "image" a set of signals which, duly processed, form an image on a screen. Besides, one can call "overlay" an approach allowing to superimpose images, whereas the term "tracking" will designate the part of the program which tracks an organ and a tool as so they are in the field **10**, to be able to determine distance separating them.

On the diagram of Figure 5a, software is structured in classes forming, by a reference link, the application that is, itself, a class and has reference **40**. Each class consists of operations and attributes. On this diagram, one can see an acquisition class **41**, which is described in a more precise way in reference to Figure 6, as well as display class **42**, conversion class **44**, image class **46**, user interface class **48**, "overlay" class **50**, "relative overlay" class **52** and "absolute overlay" **54**, stereo class **56**, "organ tracking" class **58**, and "tool tracking" class **60**. The stereo class **56** and the "tool tracking" class **60** are represented in a more detailed way, respectively, in Figures 7 and 8.

In some of the classes mentioned above, it is possible to adjust, in an automatic or voluntary way, one or several parameters that are contained in the attributes, what confers on the device a great flexibility of use.

Various attributes and operations contained in the classes are formulated in Java language, in the above-cited notation UML.

The acquisition class **41** contains operation called "newImages (right : Image, left : Image)" that permits processing of images coming from cameras **18a** and **18b**. This operation performs the synchronous acquisition of left and right images, corrects them and stores them. It is described in a more precise way in Figure 6.

The display class **42** contains operation "display (Input : Image)" that controls the display screen from signals coming from various classes composing the application.

The conversion class **44** contains the operations:

"RGBENY8 (Input: Image, Output : Image)" that permits conversion of a color image of format RGB in black and white image in format Y8;

"RGBenY8 (Input: Image, Output : Image)" that permits conversion of an image RGB in image HLS (Hue, Luminance, Saturation); and

"Reducing (factor: float, Input : Image, Output : Image)" of which the cited factor defines a rate of reduction of the resolution of an image.

The image class **46** contains information relative to an image. It has that attributes:

"Format: FormatImage", that contains information about the internal image format, i.e.: F;

Format Y8 (tones of gray);

Format RGB (color 16 , 24 or 36 bits); and

Format HLS (color 36 bits);

"Width: int", that permits the definition of the width of the image;

"Height: int", that permits the definition of the height of the image; and

"Table: Byte*" that acts as pointer on the pixels of the image.

Concerning formats RGB and HLS, one will refer advantageously to the publication of Fishkin, K., & San Rafael, A fast HLS-to-RGB TRANSFORM, Graphics GEMS, 1990, pages 448-449.

The user interface class **48** contains the operation "modifyParameters (newParams: SystemParameters)" which administers interaction between the user and the various entry peripherals of the system (keyboard, mouse, voice recognition, etc.). This operation modifies, on order of the operator, the parameters of stereo, of the tracking of organs, in particular the identification of points to be followed, the tracking of tools and by the absolute and relative "overlays", that is described later in references to the classes **50**, **52**, **54**, **58** and **60**.

The overlay class **50** contains operation "calculate (tools: ToolList, MarkList, Input:Image, InputStereo:Image, Output:Image)". Two subclasses, called "relative overlay" **52** and "absolute overlay" **54**, are derived from the overlay class **50**.

The overlay class **50** permits the calculation of the superimposing of images in absolute or relative mode according to chosen parameters, as it is explained below, and

more particularly, to perform the mixing of the images of type 2D, corresponding to an usual vision of the field **10** and to add to it images representative of the third dimension. In the case of the "relative overlay", distances are defined between two marks present in the field **10**, while in the "absolute overlay", distances are defined with regard to the endoscope.

Subclass "relative overlay" **52** contains the operation "calculate (tools: ToolList, marks: MarkList, Input : Image, InputStereo: Image, Ouput: Image)" for calculating the superimposition of images in relative mode according to parameters relating to it. It has for its attributes the parameters "RelOverlayParameters" that are the choice of the type and variable relative to this type. With this class, it is possible to superimpose just like the field of additional information by creating one or several hollow and virtual spheres centered on the extremity of tools, for example, and by representing the geometrical place of their intersections with one or several visible organs in the field **10** by a modification of the tone and/or the saturation of the corresponding pixels on the image. The extremity of the tool can be also provided with a virtual light directed in the continuation of this one, which modifies luminosity, tone or saturation of the field **10**. It is also possible to display the distance between a tool and a mark defined by the user by means of digital information, of cursors or quite other means, such as a gradient of color.

The class "absolute overlay" **54** contains operation "calculate (tools: ToolList, marks : MarkList, Input : Image, InputStereo: Image, Output : Image)" that permits the calculation of the superimposing images in absolute mode as defined above. It has for its attributes the parameters "AbsOverlayParameters" that are position and resolution of the "overlay" as well as gradient of colors.

The stereo class **56** is described in a more detailed way with reference to Figure 7. It can be noted that it contains the operations of:

"calculate (RightInput : Image, LeftInput : Image, Output : Image)" for calculating the stereo image, called also "range image", according to parameters below;

"merge (smallInput : Image, bigInput: Image, Output :Image)" for fusing the two images of disparity having the same resolution.

The class **56** has for its attributes the parameters "StereoParameters" which are the following:

Parameter of the geometry of the system for the conversion of the image of disparity in image of distance,

Parameters for the correction of the epipolar geometry,

Offsets of correction of the respective position of two images,

Space of search for disparities,

Sizes of search windows for the recognition of forms,

Reliable threshold,

Parameters for the multi-resolution (position and resolution respective various images), and

Parameters for the post-filtering.

In the present description, "recognition of forms" refers to a method for establishing the correspondences between the equivalent points of two images, left and right. This method is described in the publication of already quoted Kurt Konolige.

The "organ tracking" class **58** contains operations

"addTofollowMark (Input : Image, Mark : Position2D" and

"followMarks (Input : Image): MarkList", that allow the user to add marks **11a** to follow. It's attributes are the parameters "OrgFollowupParameters" which are the following:

number of marks to be followed;

form of the marks to be followed;

the previous position of marks to be followed; and

the current position of marks to be followed.

The "tool tracking" class **60** is described in a more detailed way in reference to Figure 8. One can note however already that it contains operation: "search (Input: Image): ToolList" having for function to search for tools in the image. It attribute parameters: "OutFollowupParameters" which are the following: maximum number of tools in the field, and types of possible tools.

In the software, the interaction of the pending classes defined above is implemented by the application of its operations, according to a sequential procedure

represented on the logical scheme of Figure 5b. The classes are represented in x-coordinate and have the same references as Figure 5a, as well as their name. Operations allowing to make application are represented in ordered, identified with the name of parameters allowing to write them over an arrow connecting application **40** with the class containing this parameter.

The represented sequence begins with the acquisition of a new image, defined by the operation "newImages (Image, Image)" contained in the acquisition class **41**. Operation "RGBenY8 (Image, Image)" converts images received in the size Y8, then operation "reduce (float, Image, Image)" allows to choose the resolution of the image. These two operations of shaping in the perspective of the subsequent treatment are contained in the class conversion **44**. The stereo class **56** allows, by the operation "calculate (Image, Image, Image)" to calculate an image of depth and, there, distances between the present various marks in the field **10**.

The following two operations, "followMarks (Image)" and "search (Image)", contained respectively in the "organ tracking" class **58** and "tool tracking" class 60, allow the tracking of marks in the field, associated respectively to organs and to tools that are there.

Operations "RGBenHLS (Image, Image)" and "calculate (ToolList, MarKList, Image, Image, Image)", contained respectively in the conversion class **44** and overlay class **50**, allow to create an image which, superimposed to the image of the field **10**, provides information relative to distances separating the marks associated to organs and to tools.

Finally, operation "display (Image)", contained in the display class **42**, transmit on the screen information allowing to see the field **10** encompassed with cameras **18a** and **18b**, as well as information relative to distances.

As can be seen in Figure 6a, the acquisition class **41** itself contains classes called optical correction **411**, filtering **412**, storage **413**, acquisition peripheral **414** and synchronisator **415**. The acquisition peripheral class **414** is connected to the acquisition class **41** through the class **415**.

The class optical correction **411** contains the operation "correct (Input: Image, Output: Image)" for correcting errors due to the optics of the endoscope. It has attribute

parameters "CorrectionParameters": Through this class, information coming from cameras **18a** and **18b** are processed in a way that their coordinates in x and y are corrected according to the optical characteristics of endoscopes, on the basis of an algorithm of correction of optical distortions described in the publication Wiliam K. Pratt, "Digital Image Processing (Second edition)", IEEE Computer Society Press Monograph, 1994.

The filtering class **412** contains the operation "register (Input: Image, Output: Image)". It has for its attributes the parameters "FilterParameters". This operation assures the treatment of images coming from cameras **18a** and **18b**, eliminates the artifacts related to the even and odd lines and corrects luminous intensity. The image obtained after the operations of these two classes are applied, is thus of appropriate quality for the continued processing.

The storage class **413** contains the operation "register (Input: Image)" which records all or any images on a digital medium, in their state before and / or after processing by the operations of the classes **411** and **412**. It is also possible to place the various states of the system in memory, the successive positions of tools and the interactions with the system of various partners, in particular the surgeon. This storage in memory allows keeping images in case of a problem. In an advantageous way, memory intended to receive them is of a permanent type, in a way that its contents cannot be modified and so can serve as a proof.

The peripheral acquisition class **414** contains the operation "to acquire (): Image". It has attributes the parameters "Peripheral" which are the following: type of signal, size of the image, and the resolution and position of the image. The video signal, initially of analog type, is transformed into digital mode by the operation contained in this class.

The synchronisator class **415** contains the operation "acquire(): Image (left: Image, right: Image)" that guarantees a synchronous acquisition of left and right images obtained after treatment by the peripheral acquisition class **414** and delivers the acquired images to the acquisition class **41**.

The diagram of the sequences of the acquisition of data is represented in Figure 6b, according to same principle as in Figure 5b. Figure 6b defines the detail of the sequence that takes place completely around the acquisition class **41**. One will note that

images resulting respectively from left and right cameras are acquired simultaneously, but processed successively. This detail is however not represented on this figure.

After the operation “newImages (Image, Image)” was effected, the acquisition class **41** effects the operation “acquire()Image (Image, Image)” and instructs the synchronisator class **415** to effect the operation “acquire ()” as for the right camera as that of the left, and simultaneously. These images are then put in memory by the operation “register (Image)” contained in the storage class **413**, then their geometry is corrected by the operation “correct (Image, Image)”. They are finally filtered by the operation “filter (Image, Image)”. Program returns then to the main loop of Figure 5b to effect the operation “RGBenY8 (Image, Image)”.

Figure 7a shows the classes derived of the stereo class **56**, and more particularly the epipolar correction class **561**, outline detection class **562**, correlation class **563** and stereo conversion class **564**.

The epipolar correction class **561** contains the operation “correct (Input: Image, Output: Image)” that transforms epipolar lines into parallel lines. It has attributes the necessary parameters “EpipGeomParams” to perform a correction of images by means of epipolar geometry, or the specific matrix of transformation in a given optics. All the useful information in that case is in the work “Determining the Epipolar Geometry and its Uncertainty: A Review “, cited above.

The outline detection class **562** contains the operation “filtrerLOG (Input: Image, Output: Image)” which filters the image by the method known under the name of LOG (Laplacian of Gaussian) described in the publication Marr, D. & Poggio, T., A computation theory of human stereo vision, Proc. Roy. Soc. Lond. B, 204, 1979. This class has for its attributes the parameters “FilterParameters” that are the coefficients of the filter LOG.

The class correlation **563** contains operation "correlate (right:Image, left:Image, Output:Image)" for defining the correspondence between the pixels of the right image and those of the left image, and to build an image of disparity. It has for its attributes the parameters "CorrelationParameters" which are:

-Offsets of correction of the respective position of the left and right images for adjusting these images so as to be able to perform correlation among them,

-Space of search for the disparity, confined on the only zone for which it is useful, for avoiding useless calculations,

-Size of search windows for the "recognition of forms", allowing also to reduce the volume of processed information, and

-Reliable threshold, which offers the possibility of defining a minimum threshold for the quality of the measures of distances.

The class stereo conversion **564** contains operation "disparityInDistance (Input:Image; Output:Image)" that transforms disparity into millimeter-length distance, by taking into account the geometry of the optics of the system. It has for its attributes the parameters "OpticalGeomParam" which are the parameters of the geometry of the optics of the device for the conversion of the image of disparity in image of distance.

Figure 7b shows the diagram of sequences for the calculation of the stereo, structured according to the same principles as for Figures 5b and 6b. In this figure, operations made on left and right images were represented in a specific way, better to differentiate operations relative to the left and right images of common operations.

Application reaches the stereo class **56** by the order "calculate (Image, Image)". Operation "correct (Image, Image)" of the class epipolar correction **561** transforms left image into epipolar coordinates. Operation "filterLog (Image, Image)" of the class detection of outlines **562** serves to define the outlines of the various objects present in the field **10** and visible on the left image. The same operations are then performed on the right image. Operation "correlate (Image, Image, Image)" of the class correlation **563** then correlates between various points of left and right images. Operation "disparityInDistance (Image, Image)" of the class stereo conversion **564** determines distances from the values of disparities. Program continues then in the main loop, such as represented to Figures 5a and 5b.

Figure 8a shows the additional classes derived of the class tracking of tools **60**. One can there see the classes detection **601**, tracking filtering **602**, tools **603** and marks **604**.

The class detection **601** makes operations "detectMarks (Input : Image): MarkList" which allows to define the position of marks in filtered images by using information relative to the geometry and to the colors of tools and marks and

"detectTools (mark: MarkList): ToolList". By this operation, program determines the various present significant points in the field **10** and representative of a tool. It corrects their adaptation and calculates the position and orientation of the tool in the space to represent it. It can furthermore regulate the position of the extremity of the tool, to be able to have a more stable image. The class detection **601** has for its attributes tools "ToolList", defined in the class **603** and which is examined later.

The class tracking filtering **602** implements operation "filter (Input : Image, Output : Image)" which allows to keep only images relative to marks and to tools but to erase the bottom. This filtering is made in size HLS. It also has attributes of tools "ToolList".

The class tools **603** contains all the information allowing to identify the various present tools in the field **10**. It has for its attributes the following parameters:

-"Type: ToolType" which contains all the information relative to tools, such as their geometry and their marking,

-"Orientation: Orientation 3D" concerning information allowing to define the orientation of the tool in the space,

-"Position: Position3D" containing information relative to the position of each of the tools in the space, and

-"Extremity: Position 3D" allowing to define the position of the extremity of tools in the space.

The consideration of the extremity of tools in a specific way allows working with a maximum of precision. It is particularly important with bistouries.

The class marks **604** has for its attributes the following parameters:

-"Center: Position 2D", providing information relative to the center of gravity of the marks present in the field **10**,

-"Surface: int" which defines the surface of each of marks,

-"Geometry: Geometry" which contains information relative to the shape of each of marks, and

-"Color: Color" in which the colors of each of marks are placed in memory.

Refering to Figure 8b, it is apparent that one reaches the class tracking of tools **60** by the order "search (Image)". The class tracking filtering **602** contains operation "filter (Image, Image)" which provides an image on having only tools and marks. The

class detection **601** contains operations "detect Marks (Image)" and "detectTools (MarkList)", then the program returns in the main loop.

As described, the device according to the invention allows surgical operations by endoscopic techniques offering a maximum of information available to be called on demand, without for all that requiring equipments hindering the work of the surgeon and his team. It is also useful in the field of the orthopedic surgery.

The same concept is not exclusively limited to the medical domain. It applies to the other numerous situations such as the inspection of channeling, the determination of the exact position of fixed or mobile objects in a difficult given space of access, etc.

It can be again noted that besides the indication of information relative to the third dimension, the other information can be associated to the displayed image, without departing from the spirit of the invention. In a not represented embodiment, relative information in third dimensions can be passed on audibly rather than by optics, the device modifying, for example, the frequency of an emitted signal, as well as providing information visually.

In another embodiment, a reticulated network is projected in infrared light on the operative field, invisible to the eye, but which can be detected by the cameras. Such a network can be used to improve the precision of measures, the number of marks being so considerably increased.